A STUDY OF SALIENCY, SEMANTIC DIFFERENTIAL SCALES AND ATTITUDES TO SCHOOL LIFE

A thesis presented to the Department of
Psychology and Sociology
University of Canterbury

In partial fulfilment of the requirements
for the degree of Master of Arts

by

COLIN JAMES CHEESMAN

May, 1967.

ACKNOWLEDGMENTS

Thanks are due to Dr. R.A.M. Gregson for supervision of this research, Dr. A.Z. Arthur for his advice in the pilot study, and Mr. D.O. Watson for technical assistance.

RESEARCH ABSTRACT

This research aimed to study the Semantic Differential Attitude Scale under conditions of high salience. Nichols and Shaw (1964) suggested that under these conditions Semantic Differential type scales were more affected by membership bias than Thurstone type scales measuring the same attitude area. The present research supports this 1964 conclusion, and suggests an explanatory hypothesis in terms of the differential effects of the various polar scales on the speed of subject penetration of scale purpose, to further understanding of the Semantic Differential scaling technique.

CONTENTS

CHAPTER ONE

AN INTRODUCTION TO THE RESEARCH PROBLEM

1. BACKGROUND TO THE PROBLEM

Over the past thirty years a considerable body of research has been concerned with social attitudes. Sociologists, social scientists and psychologists have sought to develop measures for assessing attitude through inference from verbal and non-verbal behaviour. Forming an attitude towards a social issue, makes that issue value-laden for the individual in terms of positive or negative response tendencies. Social attitudes are affected to varying degrees by group values or norms, and an individual's public behaviour often reflects group expectation rather than his own internalised values. It follows that in assessing the adequacy of an attitude measure a consideration of the effects of other influences is important. Two major aspects require careful thought: the possibility that verbal responses may differ from non-verbal behaviour in respect to the same issue; and the possibility that reported responses may have been influenced by factors other than attitude, even where there has been no conscious

effort to distort responses. Any systematic lack of correspondence between observed behaviour and reported responses can be related to three characteristics of the measurement instrument: the extent to which the purpose is clear; the extent to which the specific implications are clear; and the extent to which the responses are subject to conscious control: Cook & Selltiz, (1964). As Thurstone (1959), makes clear, what a person says and what he does, may both be inconsistent with how he feels, and it is not feasible to validate the verbal responses of an attitude scale by agreement with overt conduct. Therefore other measures are necessary to control bias.

In many situations, an opinion scale (What do you think?), and the way a respondent behaves, will be highly correlated. However, this can only be assumed in situations where we can reasonably expect people to respond to an opinion scale accurately. For example, if a scale on attitude towards religion is administered to a group of students at a church administered school, such subjects would hesitate to make known their true

opinions, if these were to deviate from expected approval of church activities. The distribution of scores on any social issue must be interpreted with due consideration for known pressures which may influence judgment. If such pressures are expected to be high, saliency effect can be expected to bias judgment towards group standards. Saliency, in this context, defines a condition which exists when the degree of importance of the stimulus material to a subject begins to affect his responses, and group pressures become important determiners of bias. Whenever test statements are rated by respondents as high-involvement questions, a state of high saliency exists, and test responses can reflect membership and reference group expectation, rather than the subject's own internalised values. It follows, that when saliency is known to be high and external pressures can be expected to affect results, it is necessary to control for this bias in the preparation of an attitude scale.

Nichols (1963), measuring attitude to college

professors with a Thurstone type scale and a Semantic Differential type scale, found very low correlations between scores obtained by these two methods. Since her subjects were college students, their attitudes to college professors were especially salient. The Nichols and Shaw (1964) study was designed to test the hypothesis that saliency affected Thurstone scores and Semantic Differential scores differently. These authors suggested that high saliency increased the tendency of the subjects to express the attitudes of the membership group, this being reflected differently by the two scales. To test this hypothesis, 'attitude to church' scales were administered to 44 church members attending church related meetings, (high saliency group), and 71 college students without church affiliations, (low saliency group); 'attitude to college professor' scales were administered to 148 college students, (high saliency group), and to 38 non college students, (low saliency group). The Pearson Product Moment correlations between the Thurstone type attitude scale and the Semantic

Differential type attitude scale for each of the four groups in this research were as follows:

Scales	High-saliency Gp.	Low-saliency Gp.	df.
Attitude to Church	.39	.76	>.01
Attitude to Professors	.29	.71	>.01

The differences between correlations for the high saliency group and the low saliency group in each of these attitude areas was significant at the .01 level. The authors concluded that the most acceptable explanation for this difference between groups was in terms of a difference in the effects of saliency on these two types of scales. They suggested that the Semantic Differential scale was more open to penetration of scale purpose, and when a high saliency condition existed, the responses to this type of scale were markedly affected by reference group pressures; penetration in this context defining the degree to which the intent of a given scale is clear to the respondents. Therefore, when both scales were sampling a highly salient attitude area, the Semantic Differential scale did not correlate highly with

the less penetrable Thurstone attitude scale.

2. THE PILOT STUDY, 1965

The results of a pilot study by this author in 1965, did not appear to support the Nichols and Shaw (1964) conclusion. The subjects for this pilot study, ranging in age from 10.6 years to 12.3 years, were from Form I classes at an Intermediate School in Christchurch, New Zealand. It was hypothesised for this study that:

(a) If differences between high-saliency and low saliency groups found to exist by Nichols and Shaw (1964) when sampling attitude through Semantic Differential and Thurstone type attitude scales, could be attributed to the differential effects of saliency bias on these two types of scales, then this effect must be the result of the degree to which the intent of each of these scales was clear to respondents.

(b) Given effective disguise of the intent of a Semantic Differential scale, the saliency effect would be minimised and this form of bias largely controlled.

To achieve disguise of a Semantic Differential attitude scale measuring attitude to school, each of the high-saliency items was surrounded by low-saliency filler items. Pre-test instructions were also designed to direct the respondents' attention away from the high-saliency items. A control group rated the same scale, but had their attention directed towards the high-saliency items in the pre-test instructions. Both groups rated a Thurstone scale which sampled the same attitude area.

Table of Correlations for a Comparison of Two Formats of a Semantic Differential Attitude Scale and a Thurstone type scale Measuring the same Attitude area

Scales	Rp	SEr	CR
Undisguised Semantic Differential Thurstone Type Scale	____ .53	.11	____ .37
Disguised Semantic Differential Thurstone type Scale	____ .59	.10	

It was considered that two possible explanations were pertinent to this failure to support the Nichols and Shaw (1964) results: either disguise had not been achieved

with the Semantic Differential scales themselves, or, for the subject population of the pilot study, differences in penetrability between the two types of attitude scales were not as marked as for the subject population of the Nichols and Shaw research.

3. THE RESEARCH FOCUS

The present research aimed to study with greater experimental control, the penetrability of the Semantic Differential type attitude scale under conditions of high salience. It was assumed for the purpose of the research that disguise had not been achieved in the pilot study (1965). This assumption was consistent with a research hypothesis that the intention of Semantic Differential type attitude scales was so easily penetrated by respondents under conditions of high saliency, that any attempt to disguise the true purpose of the scale could not be successful, and would lead to false assumptions in the evaluations of results.

CHAPTER TWO

THE ATTITUDE SCALES

1. THE SEMANTIC DIFFERENTIAL

The Semantic Differential is essentially a technique for measuring meaning through a combination of scaling method and controlled association, which was developed by Osgood et al (1952; 1957). Osgood (1952), developed a set of experimentally testable propositions as a theoretical basis for the development of the technique. He stated these propositions as: (1) A process of description or judgment can be considered as the allocation of a concept to an experimental continuum definable by a pair of polar terms; (2) that many of these ways by which meanings vary are essentially similar and can be represented on a single continuum; (3) that a limited number of these continua can define the semantic space within which the meaning of any concept will vary. Osgood utilised two methods to gather data: a judgment-time method, where intensity of association is judged by response latency to a pair of polar terms; and a graphic pencil and paper method where judgment is recorded over a seven point scale. This latter

method has since been labelled the 'Semantic Differential'.

Although the Semantic Differential was originally developed as a tool for research into the psychology of meaning, it was later recognised that the technique was applicable to a wide variety of research problems. In 'The Measurement of Meaning' (1957), Osgood, Suci and Tannenbaum have not offered it as a particular test for a restricted purpose, but as a generalised technique of measurement applicable to a wide range of purposes. Since the publication of this book investigators have continued to be interested in the Semantic Differential, and a large number of experimental applications are reported in the journals. Scott-Moss (1960), 'Current and Projected Status of the Semantic Differential', listed many varied applications of the technique as a generalised method for the measurement of meaning, and the large number of listings of research utilising this technique contained in 'Psychological Abstracts' since that date demonstrates that it has been accepted as a tool of wide utility.

2. THE SEMANTIC DIFFERENTIAL AS A TECHNIQUE TO MEASURE ATTITUDE

Osgood (1957) rates attitude measurement as one of the more significant by-products of his work in experimental semantics. He identifies 'value' with one of the major dimensions of meaning in general and offers evidence for a general principle influencing some aspects of the cognitive process - a principle of congruity.
He defines attitudes as predispositions to respond, differing from other states of readiness in that they dispose to an evaluative response. The approach - avoidance nature of these evaluative responses can be described through basic bi-polar continua with neutral reference points so that the direction and intensity of the responses become quantitatively measurable.
In terms of the Semantic Differential, the meaning of a concept is its allocation to a point in semantic space, and attitude a projection to the evaluative dimension in this space; Osgood and Suci (1955). In the preparation of an attitude scale, sets of scales with high loadings on the evaluative factor would be selected, and,

after testing, the various ratings summed over the scales to obtain an attitude score. For example, if there are M concepts and N scales, an MxN matrix would develop with scores ranging from MN to 5MN, if a five step scale is in use, where 3MN is the assumed point of least intensity of attitude. Movement away from this point in either direction would reflect bias, either positive or negative.

Osgood (1957) reports on comparative studies with Thurstone scales and Guttman scales to obtain validity and reliability coefficients for the value dimension of the Semantic Differential as a measure of attitude. The findings support a conclusion that the evaluative factor of the Semantic Differential is an index of the same dimension of meaning as is present in these other scales. The rank-order correlations with each of these attitude measures was highly significant. (p .01).

3. PRECEDENTS FOR SCALE CHOICE IN THIS RESEARCH

Tests carried out by Donahoe (1961) at four levels between first grade and college, showed that with increasing age subjects tend to agree with one another more closely in connotations of common objects.

Reporting on the acquisition of meaning for common objects in children, he was able to demonstrate a negatively accelerated curve reaching the adult asyptote at the nine year age level. Judgments on the evaluative factor were the first to reach adult norms; on this evidence the utilisation of an evaluative scaling technique in the present research would seem to be justified with an age group from eleven to thirteen. Maltz (1963), concluded that the Semantic Differential was a valid and useful instrument for the measuring of the meaning of concepts with children.

The use of a seven step scale with this instrument is not obligatory, and Maltz (1963), confirms Osgood's (1957) conclusion that primary school children seem to work better with a five-step scale. Maltz (1963) also used descriptive adjectives 'very; a little; not; a little; very;' to define the five alternatives and concluded that this verbal description improved the scale for child use. Wells and Smith (1960), investigated the relationships between four different combinations of scale formats. They presented evidence on two major

issues: (1) should scale steps be defined by adverbial qualifiers or simply left blank; (2) when several concepts are to be rated should one concept be presented at a time, or should they be presented all at once on a single sheet of paper. This research contained the following combinations of scale formats:

Multiple-stimulus / verbal;
Multiple-stimulus / non verbal;
Single-stimulus / verbal;
Single-stimulus / non verbal.

The results showed that relationships among these different scale formats were close, and gross differences in results from using the different scale formats were not to be expected. However, there was evidence of greater variability with non-verbal scales, interquartile ranges being greater than for verbal scales. Some of this variability, no doubt reflected real differences of opinion among raters, but the consistently greater variability in non-verbal ratings was a possible sign of greater random error.

Wells and Smith (1960) concluded that multiple-

stimulus-verbal scales were the best choice where distortion resulting from interaction among concepts was not likely to be troublesome. Whatever format is chosen, the choice must be made after a careful evaluation of the influences these details of format and administration may have on the results. Such an evaluation led to the choice of a single-stimulus-verbal format for the present research use of the Semantic Differential scaling technique. The choice of a multiple-stimulus scale, with the increased impact of sighting many high-saliency scales together, would have increased the probability of immediate penetration by the respondents. A single-stimulus format offered a better method for the controlled study of subject penetration of highly salient Semantic Differential attitude scales.

In 'The Measurement of Meaning' (1957) Osgood et al, in a factor analysis of meaningful judgments, found that the evaluative factor, the attitudinal variable in human thinking, accounted for a half to two-thirds of the extractable variance. Two further dimensions, a 'potency' factor, (size, weight, toughness,) and an 'activity'

factor (quickness, movement, excitement,) accounted for a large proportion of the remaining variance. As the aim of the present research was to investigate the effects of salience on revealed attitude under test conditions, three scales which have high loadings on the evaluative factor and minimal loadings on the other factors were taken as measures of attitude. These scales consisted of five-step continua each defined by a pair of polar adjectives. The scales used were:

GOOD: : : : : :BAD

HAPPY: : : : : :SAD

KIND: : : : : :CRUEL

These scales were given reversed polarity in some concept items to avoid pattern marking and to encourage an individual judgment for each item.

4. CONCEPT CHOICE FOR THE RESEARCH SCALE

Most of the studies reported in the literature using the Semantic Differential technique have required subjects to respond to single words or short phrases, and occasionally to picture stimuli. In contrast, Osgood, Ware and Morris (1961), had subjects react to complex

verbal statements, averaging a hundred words or more, as if these were unitary concepts. It does not appear that the technique need be restricted to single word stimuli, and in the present study concepts have been stated as short phrases to avoid ambiguity and provide focus.

The five concepts chosen for the Semantic Differential scale on attitude to schooling (Appendix V) were those shown by research to be most predictive in this area. Maltz (1962), after a study of the change in the meanings of concepts as a function of age, concluded that of all the concepts studied 'school' underwent the greatest quantitative change between age levels, a probable result of the fact that a child's perception of school is continually changing as a result of new experiences with this stimulus situation. Demos (1960), in a study of student ethnic groups on issues related to education, found that in all groups negative attitudes to school increased with age. These negative attitudes to school were also noted in the findings of Bosnell and Stellfire (1955), Ford (1957), and Roberts (1962). These researchers also noted that at varying achievement levels children

demonstrated differing attitudes to school tasks. Major research interest in this area of recent years has been directed towards the attitude of parents as important in shaping and influencing children's attitudes from their earliest years.

The large body of research related to school and education provides the basis for the selection of the concepts on which the Semantic Differential scales for this research are based. (Appendix I). These selected studies are not exhaustive, but serve to indicate that the choice of concepts for rating was supported by conclusions reached in experimental studies related to schooling. The filler concepts were chosen on a face validity criterion of relevance to the daily experiences of the children in the age group selected as subjects for the research. An experimental set away from the high-salience 'attitude to school' items was required in respondents as a condition of the research method. In order to plan towards this requirement, filler concepts were chosen to deal exclusively with aspects of sport considered relevant to the daily experiences of the

subject group. This battery of filler items, (Appendix V) aimed to focus the attention of respondents on a low-involvement attitude area, so as to inhibit subject-penetration of the purpose of the school attitude scales. These filler items were intended to satisfy the experimental design requirement by drawing subjects' attention away from the highly salient school-attitude items.

5. THE THURSTONE AND CHAVE SCALE USED IN THE RESEARCH

'A Study of Attitude to School Scale' used as a comparison measure in this research was devised by Fitt (1956) at the University of Auckland, New Zealand, and used as a research instrument in a study of children's attitudes to school. Items for the scale were determined according to a technique developed by Thurstone and Chave (1928), and the scale was applied to a total of 1244 Auckland school children, ranging in age from seven to eighteen years from standard one through to standard six. The test scale, (Appendix VII), has thirty items, which were originally isolated from fifty-five statements expressing different degrees of liking or disliking for school according to judges' evaluations. This scale was

considered appropriate to the ages and class of the children used as subjects in the present research.

CHAPTER THREE

EXPERIMENTAL AIMS AND METHODS

1. AIM OF THE RESEARCH

The major aim of the research was to test the generality of the Nichols and Shaw, (1964), hypothesis that under conditions of high salience Semantic Differential Attitude scales were grossly affected by bias towards functioning as a personal social acceptability scale, this effect being attributable to the relative ease with which the respondents became aware of the true purposes of these scales.

To test for the effects of salience on Semantic Differential scales, two formats of an attitude scale based on this technique were developed. A full scale was designed in two sections: part 'A' was composed of attitude-to-school items only; part 'B' of attitude-to-sport items only. By an A-B then B-A placement of the subscales, two research scales which were assumed to differ in penetrability were obtained. It was hypothesised that if saliency increases the tendency of an individual to express the attitudes of the membership group, Charters and Newcombe (1958), and if Semantic

Differential scales are especially open to penetration, Nichols and Shaw (1964), then the saliency effect should affect the validity of responses in the 'school-first, sports-last' scale where the object of an attitude (school), was highly relevant to the respondents and where no attempt had been made to disguise scale purpose. In contrast it was anticipated that the effect of the 'attitude to sport' set in the 'sport-first, school-last' scale format should be to lower subject awareness of scale purpose, or to mislead by giving the impression that 'sports attitude' was of major concern to the experimenter, thus reducing the effects of membership biases on the marking of the subsequent school items.

In order to check the relative effectiveness of the Semantic Differential scale formats in their ability to fulfil their research purposes, and to compare them with a Thurstone and Chave scale measuring the same attitude area, respondents were asked to complete a rating scale offering choice of five alternatives which stated varying degrees of scale penetration. (Results Table C).

2. DIRECT CONTROLS ON SCALE DESIGN

While preserving the essential features of the attitude scales, incidental variance was controlled by making test conditions as similar as possible for each of the three matched groups in the research. To this end scales were designed to the following criteria:

1. Instructions for the three scales were made as similar as possible within the demands of scale explanation. (Appendices IV and VI)
2. As the Semantic Differential scales were speeded, the instructions for the Thurstone scales also asked subjects to go ahead as quickly as possible. (Appendices IV and VI)
3. A practice item was necessary to explain the Semantic Differential scales; a similar item was placed at the beginning of the Thurstone scales. (Appendices IV and VI)
4. The first three items were repeated at the end of each of the Semantic Differential test scales to check subject penetration of the purposes of the scales. (Appendix V)

5. Subjects in all groups were asked to record their names before commencing, so that personal committal to attitude decision would encourage the operation of a saliency effect equally for all scales, dependent only on the relative degree of testee penetration of scale purposes.
6. All three groups were drawn from their classes at the same time, immediately after an hour-long lunch break; sources of variance due to immediate experience within each of the samples were essentially the same.
7. As the experimenter was a teacher, and as saliency could be expected to be affected by this, the classes chosen for the experiment were from a group with which he had the least contact, because of school organisation.

3. FORMATION OF GROUPS

In order to control for the possibility that the results of the Nichols (1963) research and the Nichols and Shaw (1964) research were a reflection of the peculiar subjects-scales conditions present in their experimental

procedures, subjects for this research were chosen from four separate form one classes in an Intermediate school in Christchurch, New Zealand, in contrast to the more mature subjects of the original studies. One hundred and twenty subjects were matched to three equal groups according to the following criteria:

1. To control for current teacher effect on child attitude to school, the respondents were first matched within rooms.
2. As previous research, (Appendix I), demonstrates that there are measurable sex differences in attitude to schooling at the age level of the subjects in the research, subjects were matched to sex as a second grouping criterion.
3. This was a relatively homogeneous group and age matching did not present a problem. Because of promotion policy within New Zealand schools, pupils entering form one at the Intermediate level have a relatively narrow age span.
4. The final matching criterion was an intelligence

quotient obtained through an A.C.E.R. group intelligence scale administered to each of the respondents as part of the testing of all standard four children in Christchurch schools, 1965. (Appendix III)

After matching subjects to three groups on criteria of rooms, sex, age, and intelligence rating each of these groups was given one of the research scales on the same afternoon. The two groups assigned to the two Semantic Differential scale formats were directed to the same room and given common instructions to ensure strictly comparable treatment, while the third group completed the Thurstone and Chave scale in an adjacent room. Supervisors noted that all subjects appeared to accept and mark the scales seriously, and obeyed the requirements of the standardised instructions.

CHAPTER FOUR

EXPERIMENTAL RESULTS & STATISTICAL TREATMENT

1. METHOD OF TREATMENT OF RESULTS

The graphic scale form of the Semantic Differential utilised in this research requires a subject to rate a concept in terms of a decision about its relative position on a straight line between two polar opposites. Two areas of statistical interest are apparent: firstly, the concepts which define the attitude area for judgment and determine the ostensive meanings of the various scales; secondly, the polar definitions which define another dimension of meaning by providing a means for the graphic representation of the subject's value judgments about a given concept. In the present study a decision was made to focus the research on the representational meaning dimension exemplified by the polar definitions, as this dimension could be the independent variable through which subject penetration of Semantic Differential attitude scales might best be studied. The choice of concepts was then a means for experimental control of the degree of importance of the stimulus material to the respondents.

An analysis of covariance, in which the two forms of

the Semantic Differential scale were used to study the effects of order was the principal method of statistical analysis employed in the interpretation of the results. The scales were treated as falling into three groups in terms of their polar definitions:

Good - Bad; (G.B.)

Happy - Sad; (H.S.)

Kind - Cruel; (K.C.)

In order to simplify the overall structure, and to increase the reliability of descriptive parameters, the three scales were first considered in clusters as responses to the 'sports-attitude' scales:

Cluster 1: 6 x G.B. x sports

Cluster 2: 5 x H.S. x sports

Cluster 3: 5 x K.C. x sports

where forty subjects gave a set of responses in each of the 'sports-first' items, and forty subjects in each of the 'sports-last' items. For each of the experimental groups, sixteen raw means were obtained for each of the sixteen sub-scales considered separately, and then, for

each cluster, a mean (M) and standard deviation (Sd) were derived from these raw means.

BASIC RESULTS FOR STANDARDISATION PROCEDURE

SPORTS ITEMS FIRST SCALE

	G.B.	H.S.	K.C.
Raw means: (m)	1.2 2.25 2.6 1.75 2.47 1.35	2.1 1.15 2.6 2.17 1.93	3.17 2.67 2.28 1.62 2.23
M of raw m:	1.938	1.99	2.395
Sum X^2	1.749	1.128	1.323
Sd.	.5392	.4759	.5042
Std. Scores:	$\sqrt{6}_{Sd}$ = 1.32	$\sqrt{5}_{Sd}$ = 1.064	$\sqrt{5}_{Sd}$ = 1.127
	$\frac{\sqrt{6}}{\sqrt{4}}Sd$ = .66	$\frac{\sqrt{5}}{\sqrt{4}}Sd$ = .562	$\frac{\sqrt{5}}{\sqrt{4}}Sd$ = .595

SPORTS ITEMS LAST SCALE

Raw means: (m)	1.4 1.83 1.93 2.15 1.87 1.27	1.97 1.4 2.2 1.7 1.93	2.1 2.25 1.9 1.5 2.03
M of raw m:	1.742	1.84	1.955
Sum X^2	.555	.366	.323
Sd.	.303	.270	.256
Std. Scores	$\sqrt{6}_{Sd}$ = .742	$\sqrt{5}_{Sd}$ = .594	$\sqrt{5}_{Sd}$ = .573
	$\frac{\sqrt{6}}{\sqrt{4}}Sd$ = .371	$\frac{\sqrt{5}}{\sqrt{4}}Sd$ = .256	$\frac{\sqrt{5}}{\sqrt{4}}Sd$ = .302

For each cluster, defined in terms of scale polarity, the mean of the raw mwans (M), and the standard deviation of the raw means (Sd), were used as a scaling basis for standardisation. An explanation of this standardisation procedure is set out below:

Consider the paradigm 'Happy - Sad' (H.S.):

H.S. sports = M of five sports scales

$\sqrt{5}$Sd. H.S. sports = weighted Sd. of five H.S. means

On a corresponding 'school' H.S. item, the results gave 'check-items' beginning, and 'check items' end, and four other school items. (Table D). The H.S. beginning-end check item gave two scores converted from raw scores by using H.S. sports at $\sqrt{5}$Sd. H.S. sports as a scaling basis. The four H.S. school items were pooled to give one mean H.S. score, which was standardised using H.S. sports at $\frac{\sqrt{5}}{\sqrt{4}}$ Sd. H.S. sports as a scaling basis. The same procedure was followed for G.B. ($\sqrt{6}$Sd. G.B. sports) and K.C. ($\sqrt{5}$Sd. K.C. sports), so that all scales were standardised in terms of the overall distribution parameters of their nearest covariate in terms of scale polarity. This mathematical device was used as a convenient method of

keeping the numerical values of the standard scores to the same number base in both variate and covariate for ease of statistical comparison. Using the appropriate standardisation, three parallel covariance analysis tables were set up in the standardised score form, so that each cell represented a mean score for a given subject. In each of these analyses the 'sports' score was taken as the covariate, and the 'beginning-end' criterion scores as variates. Consider the 'Good-Bad' paradigm: a two way analysis of variance with one covariate and two levels of the variate (beginning: end), the variates themselves being correlated as they are non-independent repeated observations. The analyses of variance and covariance are summarised below. Only the between subject (whole plot) comparisons have non-zero adjustments, as the covariate measure was constant for all criterion measures on the same subject.

A basic assumption underlying an analysis of covariance, is that variance due to experimental error is homogeneous within each of the treatment populations. A Cochran Test for homogeneity of variance (Results TableI) shows that the results of the test scales G.B. and H.S. are

homogeneous within acceptable limits in the Cochran formula, but that the K.C. scales are open to suspicion in this respect.

In each of the three experimental groups respondents were asked to mark over a five point scale, their individual evaluation of the attitude scale's purpose. The results of these ratings are shown in Results Table III. As two small independent examples of equal size were involved in each pairing of the scales a Kolmogorov-Smirnov test was applied to these results. (Results Tables III and IV.)

TABLE OF RESULTS

TABLE I
Cochran Test For Homogeneity Of Variance

$$C. = \frac{S^2 \text{ largest}}{S^2_j}$$

GOOD - BAD:	= $\frac{39.2575}{70.719}$	= .555	=	C
HAPPY - SAD:	= $\frac{111.529}{293.348}$	= .3802	=	C
KIND - CRUEL:	= $\frac{145.346}{208.43}$	= .6975	=	C

TABLE II
Summary Of Analysis Of Variance And Covariance
BEGINNING - END CHECK SCALES

GOOD - BAD:

Analysis Of Variance And Covariance

	X^2	XY	Y^2	$(Y)^2$
A. Diff between scales	.162	.9806	.3875	.363
Subj. W.A.	391.34	5.1409	59.28	59.21
B. Diff within scales	0	0	2.845	
A.B. Interaction effects	0	0	.0136	
Residual	0	0	.3	

Source of variance	S.S.	d.f.	Ms.	F.	P.
A. Diff between scales	.3875	1	.3875	.51	
Subj. W.A.	59.28	78	.76		
B. Diff within scales	2.845	1	2.845	11.57	$<.01$
A.B. Interaction effects	.0136	1	.0136		
Residual	.3	77	.242		
A. (Adj.)	.363	1	.363	.47	
Subj. W.A. (Adj.)	59.21	77	.769		

HAPPY - SAD:

Analysis Of Variance And Covariance

	X^2	XY	Y^2	$(Y)^2$	
A. Diff between scales	.091	.0604	.5877	.699	
Subj. W.A.	601.7193	2.432	332.93	332.93	
B. Diff within scales	0	0	6.9657		
A.B. Interaction effects	0	0	.163		
Residual	0	0	132.1837		
Source of variance	S.S.	d.f.	Ms.	F.	P.
A. Diff between scales	.5877	1	.5877	.1376	
Subj. W.A.	332.93	78	4.268		
B. Diff within scales	6.9657	1	6.9657	4.057	$<.05$
A.B. Interaction effects	.163	1	.163		
Residual	132.1837	77	1.716		
A. (Adj.)	.699	1	.699	.162	
Subj. W.A. (Adj.)	332.93	77	4.323		

KIND - CRUEL:

Analysis Of Variance And Covariance

	X^2	XY	Y^2	$(Y)^2$	
A. Diff between scales	.7391	1.764	4.2	3.991	
Subj. W.A.	628.516	37.377	106.547	104.322	
B. Diff within scales	0	0	.01		
A.B. Interaction effects	0	0	.01		
Residual	0	0	46.6		
Source of variance	S.S.	d.f.	Ms.	F.	P.
A. Diff between scales	4.2	1	4.2	3.074	
Subj. W.A.	106.547	78	1.366		
B. Diff within scales	.01	1	.01		
A.B. Interaction effects	.01	1	.01		
Residual	46.6	77	.6051		
A. (Adj.)	3.991	1	3.991	2.956	
Subj. W.A. (Adj.)	104.322	77	1.355		

The following six graphs show the standard score distributions for within scale comparisons of the variates, (School-first; School-last), and the covariate, (Sports items), for each of the polar scales 'Good-Bad', 'Happy-Sad', 'Kind-Cruel'. Graphs marked A are 'Sports First School Last' scales, and graphs marked B are 'School First Sports Last' scales.